Integrating Direct CP Search and CP-based Column Generation for the Airline Crew Assignment Problem *

Meinolf Sellmann[1], Kyriakos Zervoudakis[2],
Panagiotis Stamatopoulos[2], and Torsten Fahle[1]

[1] University of Paderborn
Department of Mathematics and Computer Science
Fürstenallee 11, D-33102 Paderborn
{sello,tef}@uni-paderborn.de
[2] University of Athens
Department of Informatics
Panepistimiopolis, 157 84 Athens, Greece
{quasi,takis}@di.uoa.gr

Abstract. We introduce a hybrid approach integrating CP-based Column Generation and pure CP methods. It is being applied to the large scale optimization problem of Airline Crew Assignment. The combination of methods from CP and OR results in an algorithm that overcomes the inherent weaknesses of each approach. First numerical results show the superiority of the hybrid algorithm in comparison to direct CP and CP-based Column Generation alone.

Keywords: Airline Crew Assignment, hybrid OR-CP method, Constrained Based Column Generation

1 Introduction

Scheduling flying crews of airline companies is a hard combinatorial problem, given the complexity of the constraints that have to be satisfied and the huge search space that has to be explored. The problem is often tackled by breaking it down into the crew pairing subproblem which consists of constructing pairings from flight legs, and the crew rostering (or assignment) subproblem which consists of assigning the constructed pairings to crew members. Although easier than the original problem, both subproblems are still hard to solve. Usually they are dealt with either by using Operations Research (OR) or Constraint Programming (CP) techniques, since they have drawn the interest of the scientific

* The presented work is supported by the PARROT Project, partially funded by the ESPRIT Programme of the Commission of the European Union as project number 24 960. The partners in the project are ILOG (F), Lufthansa Systems (D), Carmen Systems (S), Olympic Airways (GR), University of Paderborn (D), and University of Athens (GR).

communities in these two areas for many years until today. For a recent overview on optimization problems and solution techniques in the airline industry, we refer to [7, 10].

By construction, OR methods view a problem globally, taking into account all variables and usually more than one or even most constraints at a time. By calculating upper and lower bounds on the costs, they show a good ability to identify promising parts of the search space. However, they often suffer from minor local conflicts which might prevent a feasible solution from being found.

On the other hand, CP methods can efficiently handle feasibility problems by resolving local conflicts using algorithms based on arc consistency and advanced search techniques. Respectively, CP methods lack the ability to view the variables and constraints of a problem globally and therefore often have problems when stuck in local optima.

During the last decade, some work was done on the crew rostering problem. Column Generation methods have proved to be quite successful (see e.g. [3, 5, 8]). For solving the railway crew rostering problem – which is similar, but not identical to the airline crew rostering problem – Caprara et al. developed both an OR and a CLP based approach (see [1, 2]). For the latter one, a lower bound from the OR field was used to improve the algorithm.

Within the PARROT ESPRIT Project 24960 and together with our project partners, we developed a *direct* CP-based approach (DCPA) and an *indirect* approach following the CP-based column generation framework (CPCGA) (see [6]) for the Airline Crew Assignment Problem (CAP). In this paper, we show how these two approaches can be combined to overcome their inherent limitations.

2 The Airline Crew Assignment Problem

Given a set of crew members, a set of pairings, a set of rules and a cost function, a *roster* is an assignment of a subset of pairings to one specific crew member. A *schedule* is a set of rosters such that all rules are obeyed and every pairing is assigned to exactly one crew member. Rules can concern a *single crew member* or *multiple crew members*. Single crew member rules concern each individual crew member's roster, stating for example that no two temporally overlapping pairings can be assigned to the same crew member. Multiple crew member rules concern more than one crew member, stating for example that two given pairings must be assigned to two crew members out of which at least one must have a certain level of experience. The cost function associates a cost with every legal schedule and its minimization is desired.

In our case, every rule in the rule set only deals with just one single crew member, and the objective function is linear over the rosters. That means that only single crew member rules can be modeled and that the cost of the entire solution to the CAP is defined as the sum of the costs of the selected rosters. There is no more restriction to combining rosters to a solution than to obey the basic restriction that every pairing must be assigned to exactly one crew member, i.e. there are no multiple crew member rules. More formally:

2.1 Definition

Let $k, m, n \in \mathbf{N}$ and let $C := \{1, \ldots, m\}$ the *set of crew members* and $T := \{1, \ldots, n\}$ the *set of pairings*.

1. Let $R := C \times 2^T$. Every $r \in R$ is called a *roster* and R is called the *set of all possible rosters.*
2. Let $B := \{0, 1\}$ and $H := \{h_1, \ldots, h_k \mid h_i : R \to B \ \forall 1 \le i \le k\}$. Every $h \in H$ is called a *(single crew member) rule* and H is called a *rule set.*
3. A roster $r \in R$ is called *legal (with respect to a rule set H)* iff $h(r) = 1 \ \forall h \in H$. $L(H) := \{r \in R$; r is legal$\}$ is the *set of legal rosters (with respect to the rule set H).*
4. $f : R \to \mathbf{Q}^+$ is called a *cost function.*
5. The *Crew Assignment Problem (CAP)* is to minimize $\sum_{1 \le i \le m} f((c_i, t_i))$, where $(c_i, t_i) \in L(H) \ \forall 1 \le i \le m$ s.t.
   (a) $\{c_1, \ldots, c_m\} = C$
   (b) $\bigcup_{1 \le i \le m} t_i = T$ where $t_i \cap t_j \neq \emptyset \Rightarrow i = j \ \forall 1 \le i, j \le m$

Notice, that the model as stated above neither allows non linear objectives when combining rosters nor permits to restrict the combination of rosters by additional multiple crew member rules one might be interested in when tackling real life applications. Nevertheless, both DCPA and CPCGA allow to treat linear multiple crew member rules as well.

3 Two approaches to solve the CAP

In this section, we briefly introduce the two approaches for CAP we want to combine later. As our main focus in this paper is on integration, however, we do not go into too much detail here and rather refer to other papers describing the algorithms in more detail instead.

3.1 CPCGA

The definition of the CAP as stated above allows a natural decomposition into a set partitioning *master problem* and the *subproblem* of generating legal rosters. Thus, it is possible to tackle the problem with a CP-based column generation approach (see [4]). Alternately, the sub- and the master problem are being solved: Legal, individual rosters with negative reduced costs are being generated and added to the pool of rosters that have been produced so far. Then, in the master iteration step, rosters out of the pool are chosen, one for each crew member, such that each pairing is covered exactly once. The structure of the master problem is that of a set partitioning problem (SPP). It can be relaxed to a set covering problem (SCP) by only requiring the pairings to be flown by one *or more* crew members.

3.2 DCPA

The other algorithm developed to tackle the CAP is the direct CP approach. In that DCPA, each complete feasible solution of the CAP is constructed by solving the corresponding constraint satisfaction problem (see [9]). This problem is modeled by a set of variables – one variable for each pairing having as domain the set of crew members – and a set of constraints that comprise the inherent disjunctive constraint (no crew member can be assigned two overlapping pairings) and the rules in the given rule set. The problem is solved by exploring the search space, assisted by the CP propagation mechanism which prunes inconsistent possibilities as early as possible. Special heuristics that depend on the involved cost function are used to guide the search towards promising parts of the search space, as far as quality is concerned.

4 Integration

4.1 Startup Heuristic

In the CPCGA, columns are generated for each crew member sequentially. By using dual information, columns with negative reduced costs are generated, thus (when the problem is non degenerate) leading to a decrease in the continuous relaxation of the master problem. Therefore, to find high quality rosters, "good" dual values are needed. Especially in the beginning, the information contained in the dual values is very poor. This is because usually no feasible solution is known at this point and penalties stemming from dummy columns (that have to be introduced in the master problem to guarantee the existence of a solution) have a great impact on the dual values.

DCPA can help here. In the integrated approach it is used to generate a bunch of complete feasible solutions in the beginning, thereby providing one column for each crew member with every schedule found. A large variety of search methods may be used, but some of them, such as Depth-Bounded Discrepancy Search (see [11]), guarantee that columns will be adequately different from each other to make the indirect method even more efficient. Thus, a first set of columns that we know can be feasibly combined to a complete Set Partitioning Solution provides the CPCGA with the necessary "grip" to accelerate towards promising parts of the search space with respect to the "real" objective without disturbing penalties.

In any case, it is guaranteed that the indirect method finds a feasible solution that is at least as good as the best of those that were provided to it.

4.2 Combination of Columns - SPP vs. SCP

For the columns generated by the CPCGA, we find that they can be combined to Set Covering Solutions much better than to Set Partitioning Solutions (Set Covering meaning here to relax the pairing partitioning constraint, i.e. it is only required that every pairing is assigned to *at least* one crew member). That it is easier to cover all pairings than to partition the work is due to the fact that rosters are generated for one crew member at a time, first assigning work to a crew member rather than deassigning it. This is a greedy branching direction decision within the branch-and-bound roster generation process. It is motivated by the fact that dense rosters usually cost less than sparse ones.

The Set Covering Solution directs us towards a very promising region regarding total costs. The global nature of the method encourages the assumption that a big part of the suggested Set Covering Solution is part of a Set Partitioning Solution whose cost is close to that of the covering one. However, local conflicts still have to be resolved to produce a feasible Set Partitioning Solution.

Therefore, the Set Covering Solution is handed back to the DCPA which resolves local inconsistencies by deassigning and reassigning overcovered pairings (in our experiments usually less than 5%), whenever possible, or backtracking if the Set Covering Solution cannot be repaired to a Set Partitioning one. This procedure corresponds to completing a partial solution, and it is carried out using the same tools (heuristics and search methods) as in the case of generating a complete solution from scratch. When such a Set Partitioning Solution is found, it will hopefully be of better cost than the ones produced in the previous step. That solution can then be locally refined producing another set of partitioning solutions, which will then be handed back to the CPCGA and so on. This local refinement is achieved by deassigning some pairings – it is a matter of intelligent choice which ones to select –, freezing the rest and searching again for better solutions.

In the above description, the same objective function is used in both approaches. Since, however, the major burden of optimization falls on the OR part, it is also worthwhile to drop the optimization part from the CP side – partially or entirely – and to use dual values for crews and pairings provided from the OR side to produce columns that can lead to better solutions of the continuous relaxation of the master problem. Dual values can either be used as heuristic information through the search (e.g. as Shortest Path Constraint, see [4]) or even as new objective function.

The use of computationally produced information, such as dual values for guiding the CP search, is generally of great interest, since direct methods conventionally rely heavily on hand made heuristics tailored to specific problems, which limits their generic potential and applicability.

5 Numerical Results

Experiments were carried out with real life data sets of a European airline company. Figure 1 shows the costs versus time plot for CPCGA, DCPA and the consolidated approach for a data set with 7 crew members and 129 pairings. Initialy DCPA generates a solution and passes it over to CPCGA which performs one optimization iteration. The resulting solutions are passed back to DCPA which rebuilds the best solution found and then locally searches for solutions of better or equal quality containing as many rosters with negative reduced costs as possible which are then passed back to the CPCGA and the process is repeated. The same approach is used against a bigger problem as shown in figure 2. In this case however, where bigger problems are involved, more solutions are provided to the CPCGA in the initial step using a combination of locally searching areas near the first solution found with the DDS search method and a heuristic based on the original one which in addition takes care of creating diverse rosters.

The plots depict the expected behavior of CPCGA and DCPA. CPCGA steadily optimizes the objective, but the costs of the initial solution are very high low. Moreover, the time needed to find a first solution, grows with the problems size. On the other hand, DCPA finds relatively good solutions quickly by using heuristic information, but soon gets stuck. The consolidated approach benefits from both approaches: it finds good solutions quickly because of DCPA and then steadily continues to refine the solutions due to the help of CPCGA.

It can also be seen that the consolidated approach is slower than DCPA early in the experiments. It is a question of parameter tuning to find an implementation that has reasonable criteria at hand to decide whether DCPA is stuck or not.

Above the results presented here, we experiment with data from two more airline cases that differ not only in size, but also in the rules and regulations. Preliminary results show, that similar effects as concluded here can also be observed in the other airline cases.

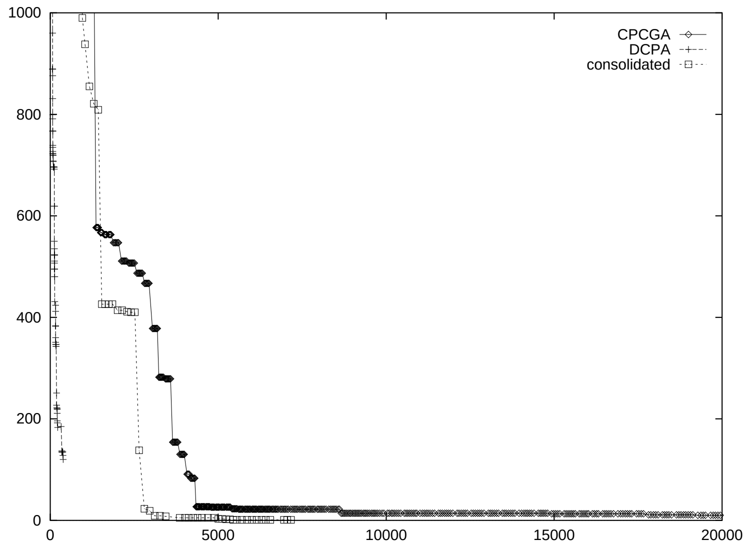

Fig. 1. Data set with 7 crew members and 129 pairings.

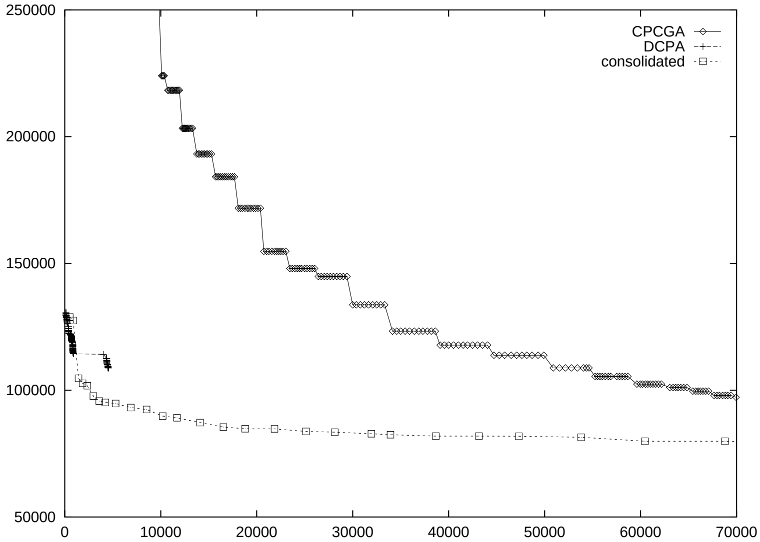

Fig. 2. Data set with 30 crew members and 279 pairings.

6 Conclusions

For the CAP as an example, we have shown how CP and OR techniques can help each other to overcome their fundamental weak points. We believe that the ideas discussed in this paper can be generalized for other problems as well, especially in connection with (CP-based) column generation. Preliminary numerical results based on data sets of a European airline show benefits of the hybrid approach proposed.

Within the PARROT project and together with our project partners, we will continue our work on the CAP and further investigate the possibilities and the potential of integration of the approaches developed.

References

1. A. Caprara, F. Focacci, E. Lamma, P. Mello, M. Milano, P. Toth, D. Vigo. Integrating constraint logic programming and operations research techniques for the crew rostering problem. *Software – Practice and Experience*, 28(1): 49–76, 1998.
2. A. Caprara, P. Toth, D. Vigo, M. Fischetti. Modeling and solving the crew rostering problem. *Operations Research*, 46(6):820–830, 1998.
3. P.R. Day, D.M. Ryan. Flight attendant rostering for short-haul airline operations. *Operations Research*, 45(5):649–661, 1997.
4. T. Fahle, U. Junker, S.E. Karisch, N. Kohl, M. Sellmann, B. Vaaben. Constraint programming based column generation for crew assignment. Technical Report tr-ri-99-212, University of Paderborn, 1999, submitted.
5. M. Gamache, F. Soumis, D. Villeneuve, J. Desrosiers, E. Gélinas. The preferential bidding system at Air Canada. *Transportation Science*, 32(3):246–255, 1998.

6. U. Junker, S.E. Karisch, N. Kohl, B. Vaaben, T. Fahle, M. Sellmann. A framework for constraint programming based column generation. *Proceedings of the CP'99*, Springer LNCS 1713:261–274, 1999.
7. R.A. Rushmeier, K.L. Hoffman, M. Padberg. Recent advances in exact optimization of airline scheduling problems. Technical Report, George Mason University, 1995.
8. D.M. Ryan. The solution of massive generalized set partitioning problems in aircrew rostering. *Journal of the Operational Research Society*, 43(5):459–467, 1992.
9. P. Stamatopoulos, G. Boukeas, K. Zervoudakis, V. Stoumpos, C. Halatsis. Parallel CP-based direct crew rostering. PARROT Deliverable D-TEC2.1, University of Athens, University of Paderborn, 1999.
10. G. Yu (editor). Operations Research in the Airline Industry. *International Series in Operations Research and Management Science*, Vol. 9, Kluwer Academic Publishers, 1998.
11. T. Walsh. Depth-bounded discrepancy search. *Proceedings of the IJCAI'97*, 1997.